

# As easy to memorize as they are to classify: The 5–4 categories and the category advantage

MARK BLAIR and DON HOMA  
*Arizona State University, Tempe, Arizona*

Recently, it has been suggested that some categories commonly used in category learning research are eliciting primarily item-level memorization strategies. A new measure of generalization, the category advantage, was introduced and used to test performance on the popular “5–4” categories. To estimate a category advantage, performance on a standard category learning task is compared with performance in an identification task, where participants learn a unique response to each stimulus. Once corrected for differences in chance expectancy, the advantage shown for the category learning task represents the degree to which participants capitalize on the natural similarity structure of the categories. In Experiment 1, the category advantage measure was validated on structured and unstructured categories. In Experiments 2 and 3, the 5–4 categories failed to produce a category advantage when tested with either of two stimulus types, suggesting that these categories elicit predominantly memorization.

Categories are great simplifiers. Behaviors can be appropriately applied to classes of objects without having to go to the trouble of learning to distinguish all the members of a class from each other. Common features of a category are experienced with each exposure to a category member, but features unique to a particular member are encountered only with exposure to that individual object. With some real-world categories, such as *birds*, the bulk of the members are never seen twice. Simple memorization cannot account for people’s ability to learn mostly nonrepeating categories; they must have some mechanism to generalize.

Models of categorization are models of generalization based on previous experience. Different theories posit different sources for that generalization. In prototype models, an average is abstracted from experience with category members. Stimuli are classified according to their distance to this average (Smith & Minda, 1998). In exemplar models, stimuli are compared with the memory traces of all the individual experienced category members (Nosofsky, 1986). Selective attention is usually incorporated in these models, allowing people to focus on the category-level regularities and ignore irrelevant information (Minda & Smith, 2001; Nosofsky, 1988). Rule-based models propose that humans develop various types of rules to guide generalization to novel instances (Erickson & Kruschke, 1998; Johansen & Palmeri, 2002;

Nosofsky, Palmeri, & McKinley, 1994). Decision boundary theories propose that humans learn boundaries that divide perceptual space into different response regions (Ashby & Maddox, 1993). All these potential bases for generalization are ways in which humans might capitalize on the natural similarity structure present in stimuli.

Whatever its underlying mechanisms, generalization allows one to use information about learned category members to improve performance on other members of the category. If one were to compare performance between people learning a pair of structured categories and people memorizing unique responses for each of the category members, one would find a performance advantage for the group learning categories (once differences in chance expectancy are taken into account). This is because the similarity between members of a structured category provides useful cues for how they should be categorized. In the memorization task, the similarity structure of the categories is irrelevant to the task because every member is paired with a unique response. We will call this natural performance advantage for learning structured categories as categories, rather than learning them individually, the *category advantage*.

Researchers have raised the possibility that participants learning categories based on binary-valued dimensions (BVD) may simply memorize each member, rather than generalize across members (Blair & Homa, 2001; Smith & Minda, 2000). BVD category instances vary on a small number of dimensions, and so these categories have severe limitations on both their sizes and structures (Smith & Minda, 2000). The most prominent BVD categories are certainly the “5–4” categories, which were introduced by Medin and Schaffer (1978). This stimulus structure has been used in at least 13 published studies and has influenced nearly a quarter century of category learning research (Anderson & Betz, 2001; Johansen &

---

Portions of the data contained in this manuscript were presented at the Annual Meeting of the Psychonomic Society in Orlando, Florida, November 2001. We thank Susan Somerville, Stephen Goldinger, Thomas Palmeri, Mark Johansen, Neil Stewart, Michael Roberts, Woodrow Clantek, and several anonymous reviewers for helpful comments on earlier versions of this manuscript. We also thank Paul Minda for giving us bugs. Correspondence concerning this article should be addressed to M. Blair, who is now at Indiana University, Department of Psychology, 1101 E. 10th Street, Bloomington, IN 47405 (e-mail: mrblair@indiana.edu).

Palmeri, 2002; Lamberts, 1995, 2000; Medin, Altom, & Murphy, 1984; Medin, Dewey, & Murphy, 1983; Medin & Schaffer, 1978; Medin & Smith, 1981; Minda & Smith, 2002; Nosofsky, Kruschke, & McKinley, 1992; Nosofsky et al., 1994; Palmeri & Nosofsky, 1995; Smith & Minda, 2000). The 5–4 categories have one 5-member and one 4-member category. Each stimulus varies on four different binary-valued dimensions (the precise structure is shown in the Appendix). These categories are more structured than random (i.e., the within-category similarity exceeds the between-category similarity) and are linearly separable. The specific stimulus type used to instantiate the four dimensions has varied across studies; geometric shapes, faces, rocket ships, and bugs are some examples. There is some suggestive evidence that participants who learn these structures are memorizing. Medin et al. (1983), using yearbook photographs to instantiate the 5–4 categories, found that it was easier to memorize unique labels than it was to learn the categories. Memorization was perhaps more likely in that study than in the usual 5–4 experiment because each stimulus provided unique information that may have made them easier to memorize. Regarding the more typical 5–4 stimulus sets, Smith and Minda (2000) pointed out that they lead to a large advantage for training stimuli over transfer stimuli. The steep drop-off in performance to new stimuli is suggestive of memorization.

There is, however, evidence that small BVD categories can elicit generalization during learning. Shepard, Hovland, and Jenkins (1961) showed that categories that are unstructured, in the sense that no feature is associated with one category more than another, are learned slower than categories that have diagnostic features. They found that the learning rate for unstructured categories was well predicted by the memorization rate for the stimuli. More structured categories, on the other hand, were learned much faster than was predicted. In a paradigm wherein participants learned both category-level and item-specific responses to each stimulus, Reed (1978) showed that for structured categories, but not for unstructured ones, category-level responses were learned faster than item-specific responses.

To give us a way of assessing the degree to which the 5–4 categories elicit memorization, we developed a method for measuring the category advantage afforded by a set of categories. The category advantage measure gives us a way of assessing the extent to which participants are learning faster than one would predict from simple memorization. Assessing the category advantage requires comparison of the data from two tasks: category learning and identification. The main obstacle to such a comparison is the difference in chance expectancy for the two tasks. In a category learning task, each stimulus is associated with one of two categories; in an identification task, each stimulus has its own unique response. The probability of guessing the correct answer is different across tasks. Moreover, in identification, this probability changes as the pool of unlearned responses diminishes. Despite these differences, it is possible to generate an expected relationship between the two tasks. The proportion of correct responses a participant will give in a particular trial block can be estimated by adding the number of learned stimuli to the number of remaining correct responses that would be expected by chance, then dividing that number by the total number of stimuli in a trial block. For example, assume that a participant has memorized the responses for five of nine stimuli. In an identification task, he/she will correctly guess one of the remaining four on average; in a categorization task, he/she will correctly guess two, yielding .67 and .78 proportion correct, respectively. The results of these calculations for a nine-stimulus set are shown in Table 1.

These expected values can be plotted on a graph, with proportion correct for a trial block of identification on the *x*-axis, and proportion correct for categorization on the *y*-axis. The resulting line represents the anticipated performance for each task as learning progresses, assuming stimulus memorization and chance guessing. However, if participants capitalize on the natural similarity structure in the stimulus set, their performance in categorization will be better than chance guessing, and their data will be above the derived line.

In order to minimize distortions due to ceiling effects, the category advantage measure, as a mean across trial blocks, should be calculated only with trial blocks that re-

**Table 1**  
Mean Proportion Correct in Identification and Categorization Tasks Given the Number of Stimuli Memorized and Chance Expectancy

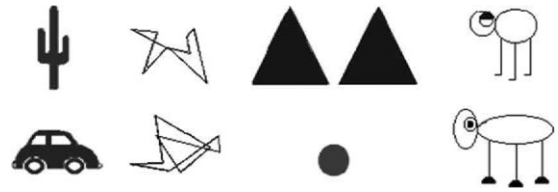
Number Memorized	Mean Number Guessed Correctly in ID Task	Mean Number Guessed Correctly in CAT Task	Total Proportion Correct in ID Task	Total Proportion Correct in CAT Task
0	1	4.5	.11	.50
1	1	4.0	.22	.56
2	1	3.5	.33	.61
3	1	3.0	.44	.67
4	1	2.5	.56	.72
5	1	2.0	.67	.78
6	1	1.5	.78	.83
7	1	1.0	.89	.89
8	1	0.5	1.00	.94

sult in less than 100% correct. Including trials in which identification performance is perfect can lead to an inflated estimate of the category advantage. In the case with nine stimuli, for example, categorization performance is predicted to be 94% when identification is at 100%. Obviously, performance can still improve on the categorization task. Any such trial blocks would be falsely counted as a categorization advantage (e.g., trial blocks where performance was perfect for both tasks would show a 6% category advantage). Including trials in which categorization performance is perfect leads to the opposite problem; the category advantage is compressed because participants in the categorization condition are at ceiling. The category advantage measure would get smaller the longer it took participants to memorize the stimuli. This is because the proportion of trial blocks with a falsely reduced category advantage would take a larger and larger proportion of the trial blocks. This distortion could be large if just a few participants had trouble learning the last few stimuli. Using only below perfect trial blocks for each condition is a simple and effective way to eliminate these distortions.

Thus far, we have introduced the idea of the category advantage and described a method for quantitatively measuring it. The primary goal of the present experiments was to assess the category advantage shown by participants learning the 5–4 categories. Because our approach is new, we first empirically verified its validity in Experiment 1, using both highly structured and unstructured categories. With that accomplished, Experiments 2 and 3 measured the category advantage yielded by the 5–4 categories. Previous work suggested that performance differences across different training stimuli and across different participants provided reason to doubt that participants were memorizing (Johansen & Palmeri, 2002; Medin et al., 1983; Nosofsky et al., 1994). In light of this work, individual differences were examined in Experiment 3.

## EXPERIMENT 1

Experiment 1 measured the category advantage for two stimulus sets, one that should and one that should not provide a category advantage. To create a stimulus set that should not provide a category advantage, we used stimuli that were simple icons, each from a different basic level category. Two examples are shown in Figure 1 (first column). These stimuli do not form real categories, and there is no similarity structure that participants might use to aid performance. Each stimulus has to be associated with the correct category by memorization. For the second stimulus set, we chose two moderately structured, ill-defined categories of random-dot polygons (Homa, 1978). Examples of these stimuli are shown in Figure 1 (second column). Blair, Homa, and Johnson (2002) used similar stimuli to study people's ability to learn categories in which instances never repeated, finding that participants easily learned these categories, even though they received novel instances in every trial. Because memorization could not account for performance,



**Figure 1.** Examples of stimuli: unique category icons and random-dot polygons (Experiment 1), geometric forms (Experiment 2), and schematic bugs (Experiment 3). In the actual experiments, the geometric-forms stimuli were either red or blue.

the participants must have been generalizing. If our measure of the category advantage is valid, we should obtain a large category advantage for the random-dot polygons and no category advantage for the unique category icons.

## Method

**Participants.** Undergraduates from Arizona State University participated for course credit. They were randomly assigned to one of four conditions: the memorization task ( $n = 16$ ) or the categorization task ( $n = 17$ ) using random-dot polygon stimuli, or the memorization task ( $n = 20$ ) or the categorization task ( $n = 24$ ) using the unique category icon stimuli.

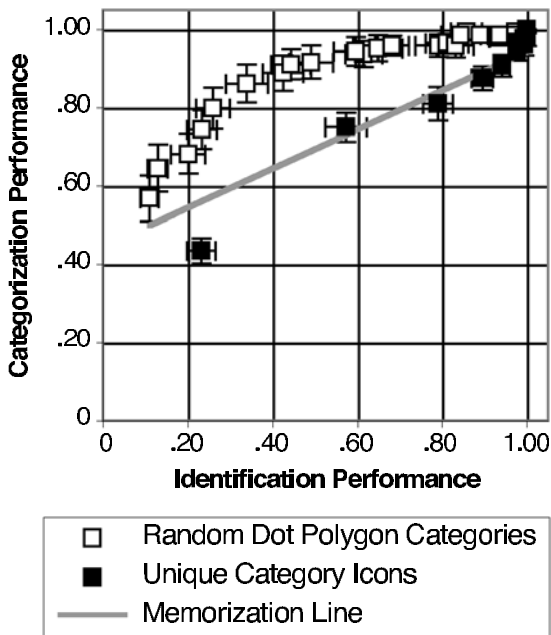
**Stimuli.** The unique category icons included the following: a suitcase, a stick figure, a baseball cap, a cactus, a car, a door, a moon, a fish, and a hammer. Each stimulus represented a different basic level category and had no obvious similarity to the others. For the categorization task, the two categories were formed by randomly assigning five stimuli to one category and four to the other. For the identification (ID) task, stimuli were randomly assigned to a unique response (1–9). The participants learned one of four different sets of random stimulus–response pairings.

For the random-dot polygon stimuli, each category had a unique prototype shape made up of nine random points on a  $50 \times 50$  grid that were connected, one with the next, creating a closed figure. Category members were created by moving each dot of the prototypes in a random direction. Varying the average distance that each point was moved changed the similarity of new polygons to the prototype. For this experiment, points were moved an average of 3.00 Euclidean units, creating a medium level of distortions (Homa & Vosburgh, 1976). Because these stimuli were complex (nine points in two dimensions each), the stochastic generation process could produce a large number of unique stimuli with little chance of repetition.

**Procedure.** In the ID tasks, each of the nine stimuli had its own unique response, one of the numerals 1–9. In the category learning tasks, participants placed the stimuli into two groups, designated by the labels “1” and “0.” For each trial, the participants saw one stimulus on a computer screen, chose a number using the keyboard, and received the correct category as feedback. Feedback was shown—along with the stimulus—for 1 sec after the response. A trial block consisted of the presentation of all nine stimuli in a random order. The experiment continued for 25 trial blocks or until the participant performed perfectly on 2 consecutive trial blocks. Participants who reached the learning criterion were scored as performing perfectly for all unfinished trial blocks.

## Results and Discussion

The data are summarized in Figure 2 with learning data from categorization and identification for each of the two stimulus types, yielding two learning curves. Each data point represents the proportion of correct responses



**Figure 2.** Experiment 1 data: Proportion correct on a particular trial block for the identification tasks (x-axis) and the categorization tasks (y-axis) by stimulus type. The memorization line reflects expected performance, correcting for chance and assuming equivalent learning rates—that is, no category advantage. Error bars reflect the standard error of the mean.

within a single trial block for both the identification and categorization tasks. Expected performance assuming equal learning rates in the identification and categorization tasks is plotted as the *memorization line*. Quantitatively, the category advantage was measured as the difference between the obtained percent correct on the categorization task minus the expected percent correct, assuming memorization, excluding the blocks with perfect performance on either task.

For the unique category icons, 10 trial blocks resulted in less than perfect performance. The mean category advantage for these trial blocks was .001, which is not significantly different from zero ( $t < 1$ ). Two-tailed tests and alphas of .05 were used for all analyses. The lack of a useful similarity structure for this stimulus set is reflected in our measure as a lack of category advantage. For the random-dot polygon stimuli, all 25 trial blocks were below perfect. The mean category advantage for these trial blocks was .145, which was significantly different from zero [ $t(24) = 11.25, p < .0001$ ]. This stimulus set yielded a sizable category advantage.

Overall, this pattern of results provides empirical support for our measure of the category advantage. We obtained a category advantage with a stimulus set known to produce generalization, and we failed to obtain a category advantage from a stimulus set designed to have no useful source of generalization. With reasonable confidence in our measure, we next tested the prevalence of generalization in the 5–4 categories.

## EXPERIMENT 2

In Experiment 2, we directly compared performance in an identification task with performance in a categorization task, now using the 5–4 categories (Medin & Schaffer, 1978) to derive an estimate of the category advantage that participants learning these categories show. Three between-subjects conditions were used: the 5–4 stimulus set in a category learning task (5–4), the 5–4 stimulus set in an ID task, and a categorization task using categories with a single diagnostic dimension (1-DIM). The maximum category advantage for stimuli similar to those used in the 5–4 can be estimated empirically by testing similar stimuli arranged into categories for which the value of one of the dimensions is perfectly associated with each category. This structure is equivalent to the Type I category structure of Shepard et al. (1961), which showed the best performance. The minimum category advantage, what we call the memorization line, is the result of a chance correction across tasks and represents equivalence of categorization and identification shown in the Type VI category structure of Shepard et al.

### Method

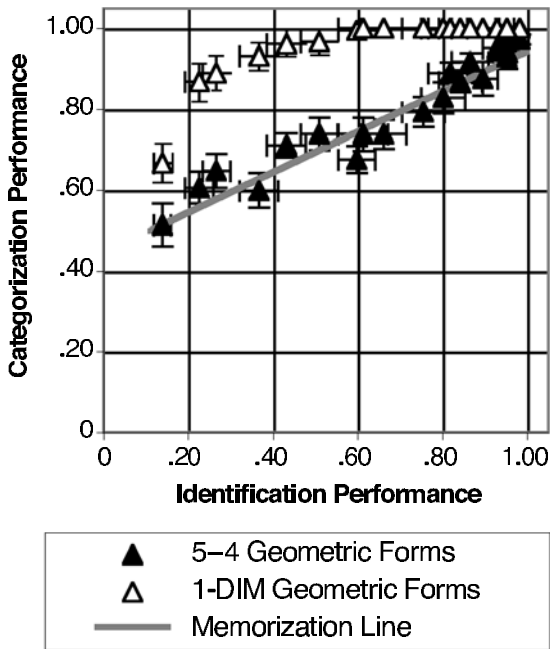
**Participants.** Undergraduates from Arizona State University participated for course credit. Participants were randomly assigned to one of three conditions: the ID task ( $n = 19$ ), the 5–4 category learning task ( $n = 16$ ), or the 1-DIM category learning task ( $n = 21$ ).

**Stimuli.** The stimuli used for Experiment 1 were geometric forms that varied in size (big or small), color (red or blue), shape (triangle or circle), and number (one or two). Examples of these stimuli are shown in Figure 1. The structures of the categories used in the 1-DIM and the 5–4 task are shown in the Appendix. The stimuli from the 5–4 condition were also used in the ID condition.

**Procedure.** The procedure was identical to that of Experiment 1, except that there were 20 trial blocks.

### Results and Discussion

The data are shown in Figure 3. From examining the 5–4 and ID data points, represented as filled triangles, it is clear that these data conform to the expected memorization relationship: Learning proceeded just as rapidly in the ID condition as in the 5–4 condition. For the 5–4 condition, perfect performance was not achieved in any trial block. The category advantage averaged across the 20 trial blocks was .016, which was at best marginally significant [ $t(19) = 1.99, p > .05$ ]. On closer inspection, only 13 of the 20 trial blocks revealed a category advantage ( $z = 1.34, p > .10$ ), and of the first 12 trial blocks (where the bulk of learning occurred for both the category and the identification task), exactly 6 values were above the memorization line and 6 were below it. Performance was much better in the 1-DIM condition; participants achieved perfect performance on the 8th trial block. The mean category advantage was .276, which was significantly greater than zero [ $t(6) = 12.80, p < .0001$ ]. The 1-DIM condition showed a category advantage for each of the 20 trial blocks, which a binomial test revealed was different from chance ( $p < .001$ ). Although the 5–4 is a relatively weakly structured category set, it is not unstructured, and therefore it could have pro-



**Figure 3.** Experiment 2 data: Proportion correct on a particular trial block for the identification task (x-axis) and each of the two categorization tasks (y-axis). The memorization line reflects expected performance, correcting for chance and assuming equivalent learning rates—that is, no category advantage. Error bars reflect the standard error of the mean.

vided some advantage over simple memorization. The structure of the categories was apparently not strong enough for participants to use, so they appear to have been memorizing. In contrast to the 5–4 categories, the well-defined categories in the 1-DIM condition yielded a large category advantage.

### EXPERIMENT 3

To explore the generality of the results of Experiment 2, we used different stimuli to instantiate the 5–4 categories. Instead of geometric forms, the stimuli were schematic bugs that varied in their type of head, body, eye, and feet (modeled after Smith & Minda, 1998). Two example stimuli can be seen in Figure 1. This stimulus type has been found to yield relatively slow learning rates (Minda & Smith, 2002). Beyond providing evidence for the generality of our results, using a more challenging stimulus set allowed us to explore an interesting prediction. If a stimulus set provides no category advantage and processing is similar for both identification and categorization, the plotted data should follow the same learning path regardless of the learning rate. The data points for a difficult stimulus set, when graphed as in Figure 2, would be compressed, relative to the data from an easier set, owing to the smaller gains in performance with each successive trial block. Nonetheless, the data should still fall on the expected learning path, presuming that both tasks get harder to equivalent degrees. Finally,

to get data that are more stable, the sample sizes were increased in Experiment 3.

### Method

**Participants.** Undergraduates from Arizona State University participated for course credit. They were randomly assigned to the ID condition ( $n = 64$ ), the 5–4 condition ( $n = 63$ ), or the 1-DIM condition ( $n = 60$ ).

**Stimuli.** The stimuli used in the three conditions (ID, 5–4, and 1-DIM) followed the same category structure as in Experiment 2, but they were line-drawn bugs similar to those used by Smith and Minda (1998). They varied on four binary valued dimensions: head, body, eye, and feet. To defend against effects due to specific physical stimulus dimensions, rather than abstract category structures, different mappings were used. There were two different physical abstract dimension mappings. For example, Dimension 1 (as shown in the Appendix) might be head shape for one group and body shape for another. Both polarities of each physical abstract dimension mapping were used. For example, a value of 1 on Dimension 1 indicated an oval head for one group and a round head for the other.

**Procedure.** The procedure was the same as in Experiment 2, except that the learning session was extended to 25 trial blocks in order to allow participants extra time to master the tasks.

### Results and Discussion

Learning difficulties are common in 5–4 category learning, and this experiment was no exception. We defined nonlearners as those participants whose mean percent correct on the last five trial blocks was less than two standard errors greater than chance (i.e., .5 for category learning tasks and .11 for identification tasks). In the 5–4 condition, there were 15 nonlearners. In the ID condition, there were 3 nonlearners. In the 1-DIM condition, there were 6 nonlearners. Since the category advantage measure relies on a comparison of learning rates across conditions, participants who failed to learn anything were excluded from the initial comparison.

The data of participants who demonstrated improvement were averaged for each trial block. Learning occurred more slowly for the bug stimuli than for the geometric shapes from Experiment 2. Over the first 20 trial blocks, the percent correct for the ID condition was an average of 22% worse for the bug stimuli than for the geometric forms. As can be seen in Figure 4, despite the challenging stimulus set, the results replicate those of Experiment 2. A large category advantage was observed in the 1-DIM condition. The 5–4 condition, however, showed no category advantage. For the 5–4 condition, the mean advantage was  $-.002$ , which is not significantly different from zero ( $t < 1$ ). Nor were there more trial blocks with a positive category advantage (9 of 25) than one might find due to chance, according to a binomial test. For the 1-DIM condition, the mean category advantage was  $.24$ , which is significantly greater than zero [ $t(24) = 15.03, p < .0001$ ]. As before, every trial block showed a positive category advantage for the 1-DIM condition. The replication of this result, obtained with a larger sample, indicated that the result found in Experiment 2 was reliable and did not vary with the specific stimuli used to instantiate the 5–4 categories.

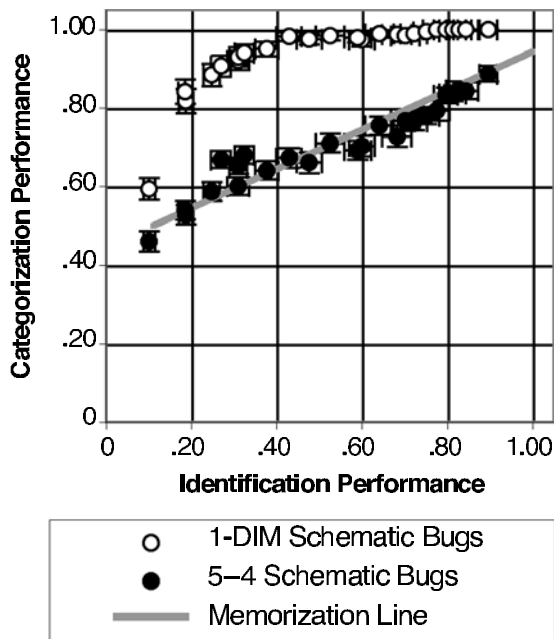


Figure 4. Experiment 3 data: Proportion correct on a particular trial block for the identification task ( $x$ -axis) and each of the two categorization tasks ( $y$ -axis). The memorization line reflects the expected performance, correcting for chance and assuming equivalent learning rates—that is, no category advantage. Error bars reflect the standard error of the mean.

The lack of a category advantage in this experiment and in Experiment 2 strongly suggests that participants who learn these categories learn them by memorization and receive no benefit from generalizing among members of the same category. When the nonlearners were included in the analysis, there was a small but significant category disadvantage of  $-.036$  for the 5–4 condition [ $t(24) = -4.00, p < .001$ ]. It seems that the successful learners are memorizers, but that not all participants are able to adopt this strategy with the bug stimulus set. This led to a replication of the findings from Experiment 2 for learners, but an increase in the number of participants who failed to learn. The nature of the stimuli (bugs vs. geometric forms) has an effect on both how easy the stimuli are to memorize, and how readily a memorization strategy is adopted, whether implicitly or explicitly. It is necessary to exclude nonlearners in order to keep the interpretation of the category advantage measure as the degree of generalization across members of the same category clean. Nonlearners cannot be interpreted as memorizing or generalizing. Indeed, the extraordinarily short reaction times of many of the nonlearners in the present study make it clear that they were not even attempting to learn at all. Although nonlearners are not profitably included in the category advantage measure, they represent an important aspect of the data from BVD category learning tasks. Although no formal model addresses the learning problems com-

monly associated with BVD categories, Murphy (2002, pp. 180–209) provides an interesting discussion of the issue.

#### Individual Stimulus and Participant Differences

It is possible that, by aggregating data across stimuli and individuals, we masked important differences that would allow us to reject the hypothesis that the 5–4 task primarily elicits memorization (Medin et al., 1983; Palmeri & Nosofsky, 1995). Both similarity-based and rule-based accounts of category learning would predict differences in performance for stimuli typical of the category or rule. In the following analyses, the data from Experiment 3 are examined for evidence of these differences.

If participants are memorizing the stimuli in both the 5–4 condition and the ID condition, the performance differences across stimuli should be similar. If participants in the 5–4 condition are showing performance differences due to the category typicality or similarity to other category members, then stimulus differences should not be related to those found in the ID condition, in which no category was learned. To examine these possibilities, the number of errors on each stimulus in the ID condition was used to predict the number of errors in the 5–4 condition in a linear regression. This analysis showed that nearly two thirds of the variance in the 5–4 condition differences could be accounted for by the ID differences, ( $R^2 = .63$ ) [ $F(1,7) = 11.98, p = .01$ ]. Since there are no categories in the ID condition, the variance accounted for by the ID data cannot be due to sensitivity to category structure. Therefore, differences in stimulus learning rates cannot be taken as strong evidence that the 5–4 category structure elicits anything more than paired associates learning. If there is an impact of category structure in the 5–4 condition, it appears to be limited, not showing itself clearly in either the category advantage measure or in the pattern of individual stimulus errors.

Some researchers have suggested that participants rely on rules to learn the 5–4 and other BVD categories (Johansen & Palmeri, 2002; Nosofsky et al., 1994). In the present experiment, individual participants' data were assessed for possible rule use by counting the number of participants with more errors on either the Dimension 1 exceptions (stimuli a5 and b6) or the Dimension 3 exceptions (a4 and b7) than on any other training stimuli. This method would detect specific participants who used a "rule-plus-exception" strategy for even a small number of trial blocks. By this criterion, 2 participants used a Dimension 1 rule and 9 participants used a Dimension 3 rule, representing a total of 23% of the participants from the 5–4 condition. In the ID condition, only 1 participant met the Dimension 1 criterion and none met the criterion for Dimension 3, a total of 2% of the participants. Because there are no categories in the ID condition, a rule-plus-exception strategy cannot apply, so this percentage can be seen as an estimate of the chance probability that a participant memorizing stimuli will end up with an

error pattern that resembles rule use. This analysis suggests that only about 20% of the participants used rules during some portion of the experiment.

A rough estimate of the extent of rule use for these participants can be gained by investigating the extent to which the exceptions have more errors than the normal stimuli. If a participant used a consistent rule-plus-exception strategy throughout the experiment, the normal patterns would have few errors, and the exceptions would have many errors. We defined four degrees of rule use—*none*, *light*, *moderate*, and *heavy*—and noted the number of participants in each group for the 5–4 condition. Participants in the none condition did not commit more errors to the exceptions. Participants in the light group committed more errors on the exceptions, but less than 50% more. Participants in the moderate group committed 50%–100% more errors on the exceptions. Participants in the heavy group committed at least twice as many errors to the easiest exception stimulus than to the hardest normal stimulus. As indicated before, the vast majority of participants did not show evidence of simple rule use. We found 5 light, 4 moderate, and only 2 heavy rule users. Aggregating the data had correctly characterized the performance of most of the people tested, but it had also masked this difference in participants' strategies. Although the present research supports the claim that some participants use rules to help them learn these categories, this strategy was relatively rare.

Some work using the 5–4 category structure and other BVD structures has adopted a rule-and-exemplar description of how participants learn these categories (Anderson & Betz, 2001; Johansen & Palmeri, 2002; Nosofsky et al., 1994). Evidence for rule use in these studies comes primarily from generalizations to new stimuli. However, use of transfer tests may provide an inflated estimate of rule use. A recent study on strategy transitions in classification provides support for this position. Using a two-category classification task, Bourne, Healy, Parker, and Rickard (1999) compared strategy use for learning and transfer. Like Johansen and Palmeri (2002), Bourne et al. found more rule use than exemplar use in early performance. Most important, though, they found that there was a large increase in rule use during the transfer task (by as much as 30%), even when an exemplar-based strategy was dominant during training. The extent to which learning and transfer strategies can differ requires further study. In light of Bourne et al.'s results, inferring learning processes from transfer performance may need to be done with more caution than previously supposed. Indeed, one strength of the category advantage method of assessing generalization is that it does not rely on the assumption that training and transfer processes are identical.

## GENERAL DISCUSSION

In the present study, we have introduced a new method of measuring generalization during category learning, quantifying the learning benefit that arises from the natural similarity structure present in categories. This ben-

efit, which we term the category advantage, was calculated by comparing identification and categorization performance on common sets of stimuli and correcting for chance expectancy. In Experiment 1, categories with a strong similarity structure showed a large category advantage and categories with no similarity structure showed no category advantage, validating our measure. In Experiment 2, the 5–4 categories (Medin & Schaffer, 1978) afforded no category advantage. That is, they were as easy to memorize as they were to classify. This result was replicated in Experiment 3 with a new stimulus set. Although an examination of individual differences revealed that a small number of participants used rules, most of the participants simply memorized.

In light of the present result that the 5–4 categories elicit primarily memorization strategies, the interpretation of data from experiments using the 5–4 categories rests critically on the assumptions made about the nature of the relationship between memorization and categorization. Exemplar theorists, who view memorization and categorization as based on the same processes, will not see a reason for concern. For them, eliciting memorization is not fundamentally different from eliciting categorization. Prototype theorists, who view memorization and categorization as involving different processes and different representations, would not accept that one could generalize from a memorization task to human category learning. Data from previous research using the 5–4 categories has been difficult for prototype models to fit, and these data are generally seen as supporting exemplar theories. If one assumes a strong relationship between memorization and categorization, then the failure of a prototype model can be seen as a critical weakness of the theory. If one assumes that there is little or no relationship between memorization and categorization, then the 5–4 category learning task has little to do with categorization and is therefore an inappropriate test of the prototype model. Smith and Minda (2000), suspecting that the 5–4 categories elicited mostly memorization, demonstrated that a prototype model could give a good account of 5–4 data if an exemplar memorization component was added to the model.

The relationship between categorization and memorization is still a matter of intense debate. Some researchers have found dissociations between recognition and categorization (Knowlton & Squire, 1993). Others have argued that these differences can be accounted for by a single-system exemplar model (Nosofsky & Zaki, 1998). Still other researchers have argued that a prototype model with a memorization component can fit dissociation data better than an exemplar model can (Smith & Minda, 2001). Until the relationship between memorization and categorization is clarified, data from the 5–4 category learning tasks do not have a clear interpretation.

Considering the present results, it is sensible to ask how widespread memorization is in the categorization literature. Because some small, seemingly poorly structured categories can provide a category advantage (Shepard et al., 1961), it is difficult to say without further direct

measurements. There is evidence that memorization occurred in Medin and Schwanenflugel's (1981) study of the importance of linear separability in category learning. Blair and Homa (2001) hypothesized that Medin and Schwanenflugel had failed to find a linear separability constraint on category learning, because the categories that they used, which were similar to the 5–4, elicited exemplar memorization processes (also suggested by Murphy, 2002). If so, people would not be constrained to find linear solutions, unlike when they use other categorization processes (e.g., rules or prototypes) that may be constrained to a linear solution. In order to make memorization a less attractive strategy, Blair and Homa used a four-category learning task. They also manipulated category size, using both three-member categories and nine-member categories, but keeping proportion of exceptions constant. They found evidence of a linear separability constraint for some participants from both the small and large categories. Importantly, the percentages of participants who could not learn the exceptions to the rule increased from 9% for three-member categories to 81% for nine-member categories. Increasing the stimulus pool meant less memorization and more participants operating under a linear separability constraint. This pattern of results was not adequately predicted by an exemplar generalization model. Smith, Murray, and Minda (1997) have also shown that using larger, well-differentiated categories increases the importance of linear separability. The prevalence of memorization in the present study gives credence to the hypothesis that memorization was also prevalent in the categories used by Medin and Schwanenflugel.

To summarize, in this article we have introduced a new method for determining the degree to which participants generalize during category learning. This method has several positive attributes. It does not assume any particular processes underlying categorization. Because the category advantage measure is calculated using category acquisition data, it does not rely on the assumption that training and transfer processes are identical. It is simple to obtain and calculate. Finally, it can be applied to any stimulus type and any category structure, with obtained category advantage scores being comparable across vastly different categories. In addition to introducing the category advantage, this article has provided empirical evidence that the widely used and influential 5–4 categories are learned chiefly by memorization, with no generalization across category members. Because the relationship between stimulus learning and category learning is not yet clearly understood, the generalizability of data from the 5–4 categories, and perhaps related BVD categories as well, to categorization cannot be determined, thereby limiting the usefulness of these categories.

## REFERENCES

- ANDERSON, J. R., & BETZ, J. (2001). A hybrid model of categorization. *Psychonomic Bulletin & Review*, *8*, 629-647.
- ASHBY, F. G., & MADDUX, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, *37*, 372-400.
- BLAIR, M., & HOMA, D. (2001). Expanding the search for a linear separability constraint on category learning. *Memory & Cognition*, *29*, 1153-1164.
- BLAIR, M., HOMA, D., & JOHNSON, L. (2002, November). *Category learning and recognition when instances do not repeat*. Poster session presented at the 43rd Annual Meeting of the Psychonomic Society, Kansas City, MO.
- BOURNE, L. E., HEALY, A. F., PARKER, J. T., & RICKARD, T. C. (1999). The strategic basis of performance in binary classification tasks: Strategy choices and strategy transitions. *Journal of Memory & Language*, *41*, 223-252.
- ERICKSON, M. A., & KRUSCHKE, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107-140.
- HOMA, D. (1978). Abstraction of ill-defined form. *Journal of Experimental Psychology: Human Learning & Memory*, *4*, 407-416.
- HOMA, D., & VOSBURGH, R. (1976). Category breadth and the abstraction of prototypical information. *Journal of Experimental Psychology: Human Learning & Memory*, *2*, 322-330.
- JOHANSEN, M. K., & PALMERI, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology*, *45*, 482-553.
- KNOWLTON, B. J., & SQUIRE, L. R. (1993). The learning of categories—parallel brain systems for item memory and category knowledge. *Science*, *262*, 1747-1749.
- LAMBERTS, K. (1995). Categorization under time pressure. *Journal of Experimental Psychology: General*, *124*, 161-180.
- LAMBERTS, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227-260.
- MEDIN, D. L., ALTOM, M. W., & MURPHY, T. D. (1984). Given versus induced category representations—use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 333-352.
- MEDIN, D. L., DEWEY, G. I., & MURPHY, T. D. (1983). Relationships between item and category learning—evidence that abstraction is not automatic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *9*, 607-625.
- MEDIN, D. L., & SCHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207-238.
- MEDIN, D. L., & SCHWANENFLUGEL, P. J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning & Memory*, *7*, 355-368.
- MEDIN, D. L., & SMITH, E. E. (1981). Strategies and classification learning. *Journal of Experimental Psychology: Human Learning & Memory*, *7*, 241-253.
- MINDA, J. P., & SMITH, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 775-799.
- MINDA, J. P., & SMITH, J. D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 275-292.
- MURPHY, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
- NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.
- NOSOFSKY, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 700-708.
- NOSOFSKY, R. M., KRUSCHKE, J. K., & MCKINLEY, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 211-233.
- NOSOFSKY, R. M., PALMERI, T. J., & MCKINLEY, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53-79.
- NOSOFSKY, R. M., & ZAKI, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, *9*, 247-255.
- PALMERI, T. J., & NOSOFSKY, R. M. (1995). Recognition memory for ex-



- ceptions to the category rule. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 548-568.
- REED, S. K. (1978). Category vs. item learning: Implications for categorization models. *Memory & Cognition*, **6**, 612-621.
- SHEPARD, R. N., HOVLAND, C. I., & JENKINS, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, **75**(13, Whole No. 517), 42.
- SMITH, J. D., & MINDA, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1411-1436.
- SMITH, J. D., & MINDA, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 3-27.
- SMITH, J. D., & MINDA, J. P. (2001). Journey to the center of the category: The dissociation in amnesia between categorization and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 984-1002.
- SMITH, J. D., MURRAY, M. J., & MINDA, J. P. (1997). Straight talk about linear separability. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 659-680.

**APPENDIX**  
**Category Structure for the 5-4 and 1-DIM Categories**

5-4 Category Structure		1-DIM Category Structure	
Category A	Category B	Category A	Category B
1 1112	6 1122	2 121	1 121
2 1212	7 2112	2 122	1 122
3 1211	8 2221	2 211	1 211
4 1121	9 2222	2 212	1 212
5 2111		2 221	

Note—In the 1-DIM categories, the first dimension is perfectly correlated with the correct category.

(Manuscript received September 16, 2002;  
revision accepted for publication July 28, 2003.)